RESEARCH ARTICLE OPEN ACCESS

Multi Heart Disease Classification in ECG Signal Using Neural Network

Theynisha .R, Shree Soundharya .R, Umarani .S, AnithaNithya. R
Department of Computer Science and Engineering
Sri Krishna College of Technology, Coimbatore
Tamil Nadu – India

ABSTRACT
Data mining is the process of data analyzing from various perspective and combining it into useful information. This technique is used for finding various heart diseases. Based on risk factor the heart diseases can be predicted very easily. The main aim of this paper is to evaluate different classification techniques in heart diagnosis. First, dataset is to be filtered using effective non local means filter algorithm. After taking risk factor, dataset is to be classified by using K-Nearest Neighbor (KNN) and Neural Network (NN).Compared to KNN, Neural Network provides better performance. After classification, performance criteria including accuracy, precision, Fmeasure is to be calculated. The comparison measure reveals that Neural Network is the best classifier for the diagnosis of heart disease on the existing dataset.

Keywords:- Data mining, Heart diagnosis, K-Nearest Neighbour (KNN) ,Neural Networks (NN), Non loca lmeans filter (NLM).

I. INTRODUCTION

The main reason for death worldwide, including South Africa is heart attack diseases and possible detection at an earlier stage will prevent these attacks. Medical practitioners generate data with a wealth of concealed information present, and it's not used effectively for predictions. For this reason, the research converts the unused data into a dataset for shaping using different data mining techniques. People die having encountered symptoms that were not taken into considerations. There is a requirement for medical practitioners to predict heart disease before they occur in their patients. The features that increase the chances of heart attacks are smoking, lack of physical exercises, high blood pressure, high cholesterol, unhealthy diet, detrimental use of alcohol, and high sugar levels . Cardio Vascular Disease (CVD) constitutes coronary heart, cerebrovascular or Stroke, hypertensive heart disease, congenital heart, peripheral artery, rheumatic heart disease, inflammatory heart disease.

Data mining is a knowledge discovery technique to examine data and encapsulate it into useful information. The current research intends to forecast the probability of getting heart disease given patient data set. Prophechies and descriptions are principal goals of data mining; in practice Prediction in data mining involves attributes or variables in the data set to locate an unknown or future state values of other attributes. Description emphasize on discovering patterns that describes the data to be interpreted by humans.

II. PROBLEM FORMULATION

Support Vector Machine (SVM) is a category of universal feed forward networks like Radial-basis function networks, pioneered by Vapnik. SVM can be used for pattern classification and nonlinear regression. More precisely, the support vector machine is an inexact implementation of the method of structural risk minimization. This principle is based on the fact the rate of error of a learning machine on test data is bounded by the total sum of the training-error rate and term which is based on the Vapnik-Chervonenkis (VC) dimension. The support vector machine can provide good generalization performance on pattern classification problem.

2.1 Optimal Hyperplane for patterns:

Consider the training sample where xi is the input pattern for the ith instance and yi is the corresponding

target output. With pattern represented by the subset yi= +1 and the pattern represented by the

Subset yi= -1 are linearly separable. The equation in the form of a hyperplane that does the separation is

$$W^T X + b = 0 \qquad (1)$$

Where x is an input vector, w is an adjustable weight vector, and b is a bias. Thus

$$W^T X_i + b \geq 0 for y_i = +1 \qquad (2)$$

$$W^T X_i + b < 0 for y_i = -1 \qquad (3)$$

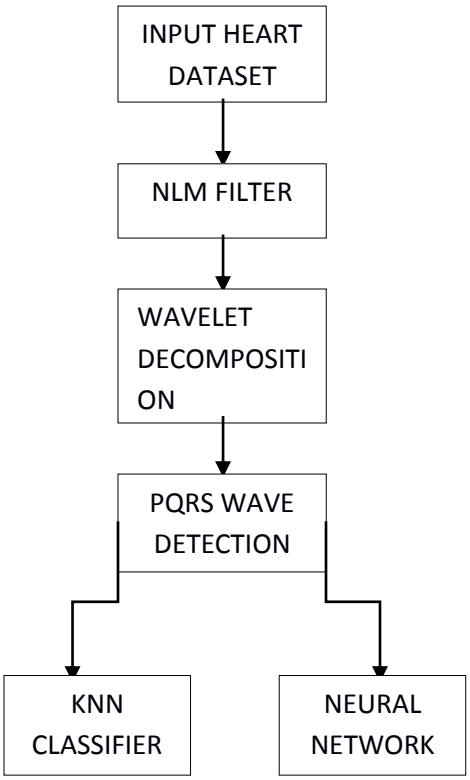
The discriminant function provides an algebraic measure of the distance from x to the optimal hyperplane for the optimum values of the weight vector and bias, respectively.

1. SVM is a binary classifier. To perform a multi-class classification, pair-wise classifications can be deployed (one class with all others, for all classes).

2. more expensive for computation, thus runs slow

III. PROPOSED FRAMEWORK

Block diagram



3.1 Non local means algorithm

- If only one iteration of the NLMeans filter is applied, the solution has generally not converged to a (local) optimum. Practically, local neighborhoods that only have few similar neighborhoods may still contain noise after one iteration (unless h is chosen large enough; but this would cause over smoothing in other regions). In this paper, we propose to keep track of the noise variance at every location in the image and to remove the remainder of the noise as a post-processing step using a local filter.

- The NLMeans algorithm can be considered to be the first iteration of the Jacobian optimization algorithm for the cost function. This would suggest that applying the NLMeans algorithm iteratively would further decrease the cost function.

- The choice of the Leclerc robust function is somewhat arbitrary. We will see that further improvements can be achieved by using the Bisquare robust function.

- The existing NLMeans algorithms assume that the image noise is white (uncorrelated), while in most practical denoising applications the noise is correlated. In this case, computation of the similarity based on the Euclidean distance is hampered which eventually leads to a poor denoising performance.

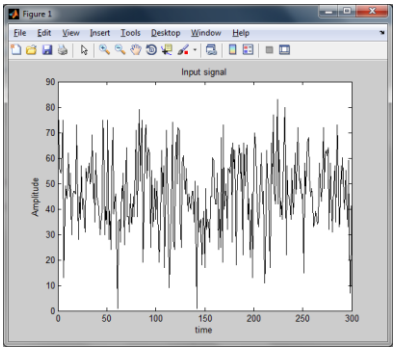


Fig 1. Input Signal

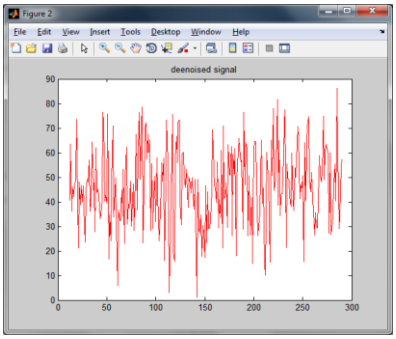


Fig 2. Denoised signal using NLM Filter

3.2 Wavelet transform

The convolution of the wavelet function ψ(t) is wavelet transform with the signal x(t). The discrete wavelets that are orthonormal dyadic are associated with scaling functions φ(t). The approximation coefficients S can be produced by convoluting the signal with the scaling function . The discrete wavelet transform (DWT) can be written as

$$T_{m,n} = \int_{-\infty}^{\infty} x(t)\varphi_{m,n}(t)dt \qquad (4)$$

By choosing an orthonormal wavelet basis ψm,n(t) the original [5] can be reconstructed . Approximately coefficient of the signal at the scale m and location n can be written as

$$S_{m,n} = \int_{-\infty}^{\infty} x(t)\phi_{m,n}(t)dt \qquad (5)$$

But the discrete input signal is of N finite length. So the range of scales that can be examined is 0 <m

$$x_0(t) = x_M(t) + \sum_{m=1}^{M} d_m(t) \qquad (6)$$

Where the mean signal approximation at scale M is xM(t) = SM,nφM,n(t) and detail approximation of signal that corresponds to scale m, for finite length signal is

$$d_m(t) = \sum_{n=0}^{M-m} T_{m,n}\psi_{m,n}(t) \qquad (7)$$

The approximation of signal at a specific scale is a integration of the approximation and detail at the next lower scale.

$$x_m(t) = x_{m-1}(t) - d_m(t) \qquad (8)$$

In the present work Daubechies [6] wavelet is chosen although the Daubechies algorithm has more complexity and complicated computations, but this algorithm picks up minute detail that is missed by other wavelet algorithms, like Haar wavelet algorithm. A signal cannot represented well by one member of the Daubechies family, that may still be efficiently represented by another.

3.3 PQRS wave detection

3.3.1 Algorithm steps

Step1: Output wavelet signal is read and the length is calculated.

Step2: The signal is decomposed using db6 wavelet.

Step 3: The detailed coefficients 3 , 4 and 5 are selected, as these coefficients have most of the energies of QRS complex .

Step 4: The wave is reconstructed using detail coefficient 3, 4, 5,(D1=d3+d4+d5)

Step 5: A function d4*(d3+d5) /2n is defined as it reduces the signal's oscillatory nature where d3, d4, d5 are the 3th , 4th ,5th detail coefficients and n is the level of decomposition.

Step6: Derivate till level 5 is made using the transfer function.

Step7: The differential equation is applied to the signal, by using the transfer function and taking the amplitude response as

y(nT)=(T/8)(-x(nT-2T)-2x(nT)+2x(nT+T)+x(nT+2t))

Step8: Square the signal point by point using the equation

y(nT)=(T/8)(-x(nT-2T)-2x(nT)+2x(nT+T)+x(nT+2t))

to emphasize R wave from the ECG signal.

Step9: A moving window is combined using the equation

Y=(1/N)*[x(nT-(N-1)T)+x(nT-(N-2)T)+x(nT)] to obtain the information of waveform feature.

Step10: Calculate the threshold value corresponding to the product of max and mean of the signal so that the end points of the moving window can be located.

Step11: The PQRST peaks are located based on the amplitude of the signal within each moving window.

Step12: Consider the positions of two consecutive same labeled peaks and store it to calculate the time intervals.

Step13: Diagnosis of various heart diseases is done by comparing ground truth conditions with the data.

Rise time = Rwave – Swave /10

Fall time = Rwave – Qwave /10

PQ interval = Pwave – Qwave/100000

ST interval = Swave – Twave /5000000000

QT interval = Qwave – Twave/500000

QRS interval = (Qwave + Rwave)-Swave/100000.

3.4 classifier

- KNN Classifier
- Neural Network

3.4.1 KNN for Classification

Let's see how to use KNN for classification. In this case, we are given few data points for training and a new unlabelled data for testing. Our aim is to set the class label for the new point. The algorithm posses different behavior based on values of k.

Case 1: k = 1 or Nearest Neighbor Rule

Label a point x. Find the point closest to x . Let it be y. Now nearest neighbor rule requires to assign the label of y to x. This is very simple and sometimes even counter intuitive. If it is felt that the procedure will result a huge error, it is right – but there is a catch. This reasoning is applicable only when the number of data points is not great.

If the number of data points is very large, then there is a very greater chance that label of x and y is same. An example might help – consider a (potentially) biased coin.Toss it for 1 million time and favours head 900,000 times. Then most likely your next call will be head.Similar argument can be used.

Try for an informal argument - Assume all points are in a D dimensional plan. The number of points is reasonably large. This implies that the density of the plane at any point has higher density. In other words, within any subspace there is sufficient number of points. Take a point x in the subspace which also has a lot of neighbors. Now consider y as the nearest neighbor. If x and y are closer, then we can assume that probability that x and y belong to similar class is relatively same – Then by decision theory,values x and y have the same class.

There is an excellent discussion in the book "Pattern Classification" written by Duda and Hart h about this Nearest Neighbor rule. One of their results is to obtain a better tight error bound to the Nearest Neighbor rule.

Case 2: k = K else k-Nearest Neighbor Rule

This is 1NN's straightforward extension. Basically what we do is that we try to find the k nearest neighbor and do a majority voting. The value k is odd when the number of classes is 2. Consider k = 5 and there are 3 instances of C1 and 2 instances of C2. Here, KNN says that new point has to label as C1 as it forms the majority.Similar argument is followed in case of multiple classes.

One of the straight forward extensions is not to give 1 vote to all the neighbors. A very basic thing to do is weighted *KNN* where each point has a weight which is calculated using its distance. For instance

under inverse distance weighting, each point posses a weight equal to the inverse of its distance to the point is classified. This means that neighboring points has a greater vote than the farther points.

It is quite obvious that the accuracy increase when you increase k but the computation cost also increases.

3.4.2 Training the Neural Network:

Here the neural network is trained with the database of Heart Diseases using feed forward neural network model and back propagation learning algorithm having momentum and variable learning rate.13 neurons are present in the input layer of the network to specify each attribute and so the database consists of 13 attributes. The class numbers are four: 0 – normal person,1- first stroke, 2- second stroke and 3- end of life. These classes are represented using two neurons in the output layer. The details of the back propagation algorithm is specified in the above is used for training the neural network . More neural networks are constructed with and without hidden layers, i.e, single layer and multi layer networks and trained using the dataset of heart diseases. Relationship between the number of epochs and the sum of squares of errors during training process for various networks can be observed .

IV. PERFORMANCE MEASURE

A group of metrics including accuracy rate (ACC), true positive (TPR), false positive rate (FPR), and positive predictive value (FPV), were employed to quantify the locating accuracy. The definition for the ACC, TPR, FPR, and FPV can be described as

ACC or Accuracy rate= (TP+TN) / (TP+TN+FP+FN)

TPR or True positive rate = TP / (TP + FN)

FPR or False positive rate = FP / (FP + TN)

PPV or Positive positive value = TP / (TP + FP)

V. EXPERIMENTAL RESULT

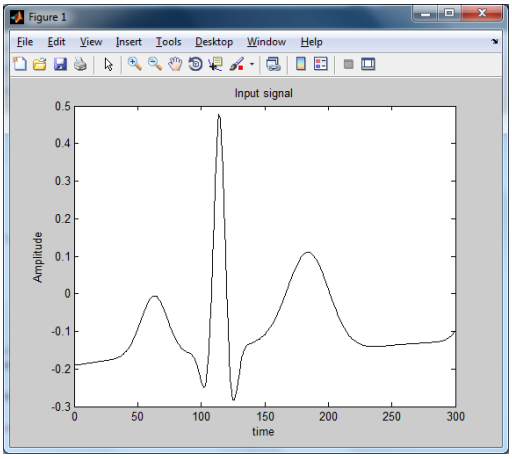


Fig 3. Input signal

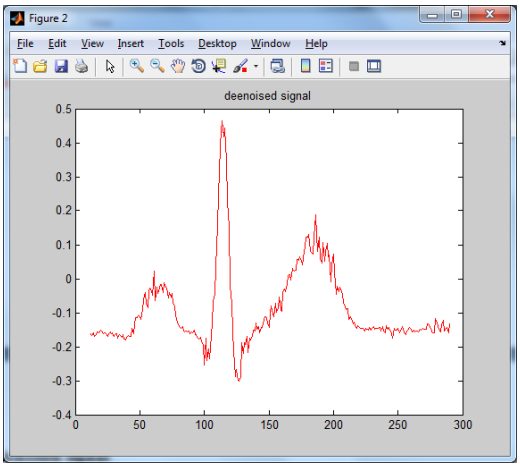


Fig 4. Denoised signal

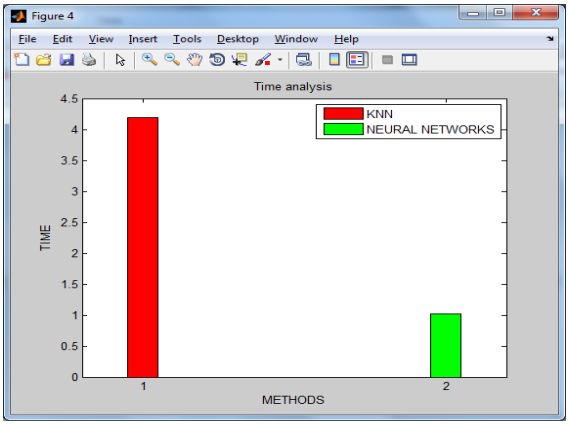


Fig 5. Time analysis graph for KNN and Neural Network

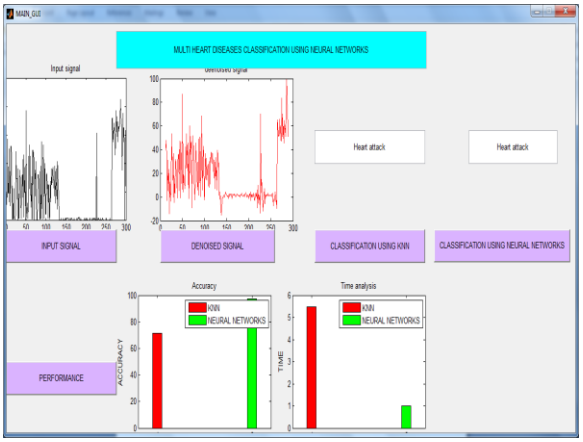


Fig 6. Heart attack

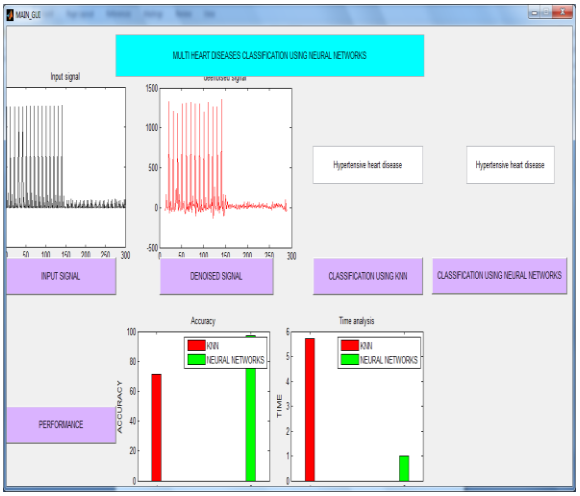


Fig 7. Hypertensive heart disease

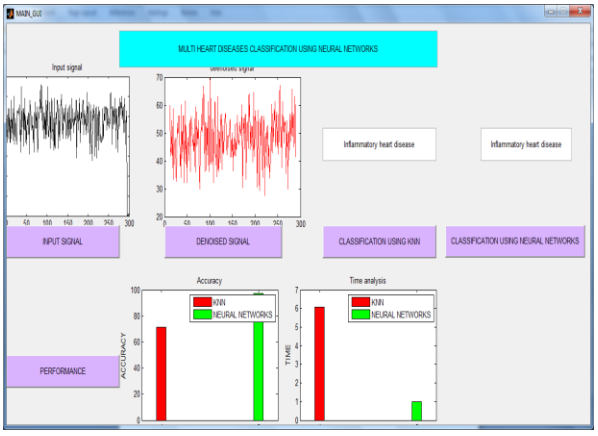


Fig 8. Inflammatory heart disease

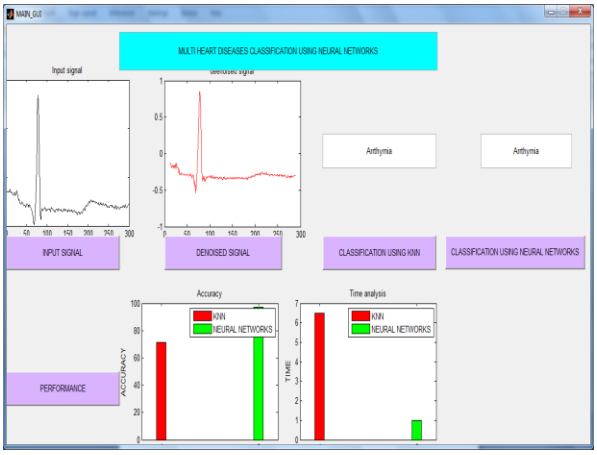


Fig 9. Arrthymia

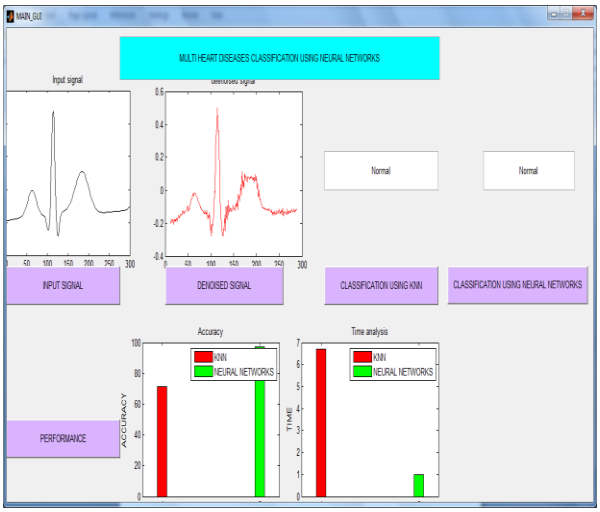


Fig 10. Normal

VI. CONCLUSION

The experiment is conducted with the dataset of Heart Disease by contemplating the single and multilayer neural network modes. Backpropogation algorithm along with momentum and variable learning rate is used to train the networks. To examine performance of the network various test data are given as input to the network. Number of Epochs Training-Blue Goal-Black International Journal of 7 neuron in all hidden and output layers to accelerate the learning process.Thus Neural networks tech provides satisfactory results for classification.

REFERENCES

[1] Amma, N.G.B "**Cardio Vascular Disease Prediction System using Genetic Algorithm**", IEEE International Conference on Computing, Communication and Applications, 2012

[2] Chaitrali S Dangare "**Improved Study Of Heart Disease Prediction System Using Data Mining Classification Techniques**", International Journal Of Computer Applications,Vol.47,No.10 (June 2012).

[3] Beshiba Wilson, Dr.Julia Punitha Malar Dhas' **A Survey of Non-Local Means based Filters for Image Denoising**'International Journal of Engineering Research & Technology ,Vol.2 - Issue 10 (October - 2013

[4] Y. Guo, Y. Wang, T. Hou, " **Speckle filtering of ultrasonic images using a modified nonlocal-based algorithm**", Biomedical Signal Processing and control, Elsevier, pp.129-138, 2011.

[5] Pierrick Coupe, Pierre Hellier, Charles Kervram and Christian Barillot, "**Bayesian non local means-based speckle filtering**",IEEE International Symposium on Biomedical Iamging,2008.

[6] Sayantan Mukhopadhyay1 , Shouvik Biswas2 , Anamitra Bardhan Roy3 , Nilanjan Dey4' **Wavelet Based QRS Complex Detection of ECG Signal**' International Journal of Engineering Research and Applications (IJERA) Vol. 2, Issue 3, May-Jun 2012, pp.2361-2365

[7] 'ECG Signal Analysis Using Wavelet Transforms', C. Saritha, V. Sukanya, Y. Narasimha Murthy, Department of Physics and Electronics, S.S.B.N. COLLEGE (Autonomous), Andhra Pradesh, India

[8] Tompkins WJ, Pan J **A real time QRS detection algorithm**, IEEE Transactions on Biomedical Engg. BME-32, No. 3:230-235, March 1985.

[9] Algorithm M.Akhil jabbar* B.L Deekshatulua Priti Chandra bInternational' **Classification of Heart Disease Using K- Nearest Neighbor and Genetic Algorithm'** Conference on Computational Intelligence: Modeling Techniques and Applications (CIMTA) 2013